

# Elimination of the Translational Kinetic Energy Contamination in pre-Born–Oppenheimer Calculations

Benjamin Simmen<sup>a</sup>, Edit Mátyus<sup>b1</sup>, Markus Reiher<sup>a2</sup>

<sup>a</sup>ETH Zurich, Laboratorium für Physikalische Chemie,  
Wolfgang-Pauli-Str. 10, 8093 Zurich, Switzerland

<sup>b</sup>Laboratory of Molecular Structure and Dynamics, Institute of Chemistry, Eötvös University  
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary

December 15, 2012

## Abstract

In this paper we present a simple strategy for the elimination of the translational kinetic energy contamination of the total energy in pre-Born–Oppenheimer calculations carried out in laboratory-fixed Cartesian coordinates (LFCCs). The simple expressions for the coordinates and the operators are thus preserved throughout the calculations, while the mathematical form and the parametrisation of the basis functions are chosen so that the translational and rotational invariances are respected. The basis functions are constructed using explicitly correlated Gaussian functions (ECGs) and the global vector representation.

First, we observe that it is not possible to parametrise the ECGs so that the system is at rest in LFCCs and at the same time the basis functions are square-integrable with a non-vanishing norm. Then, we work out a practical strategy to circumvent this problem by making use of the properties of the linear transformation between the LFCCs and translationally invariant and center-of-mass Cartesian coordinates as well as the transformation properties of the corresponding basis function parameter matrices. By exploiting these formal mathematical relationships we can identify and separate the translational contamination terms in the matrix representation of the kinetic energy operator in the LFCC formalism.

We present numerical examples for the translational contamination and its elimination for the two lowest rotational energy levels of the singlet hydrogen molecule, corresponding to para- and ortho-H<sub>2</sub>, respectively, treated as four-particle quantum systems.

---

<sup>1</sup>corresponding author; e-mail: matyus@chem.elte.hu

<sup>2</sup>corresponding author; e-mail: markus.reiher@phys.chem.ethz.ch

# 1 Introduction

Expressions for the calculation of intrinsic properties of molecules should be free of contributions from the overall translation of the system. In the commonly introduced Born–Oppenheimer (BO) approximation, the nuclei are fixed, and thus the translational contribution is automatically separated. In several combined electron-nuclear orbital approaches [1–3] the translational dependence is eliminated automatically by fixing one or a few heavy particles. Here, we consider molecules as many-particle quantum systems with electrons and nuclei both treated as quantum particles on equal footing in the pre-Born–Oppenheimer (pre-BO) quantum theory.

Traditionally, in rovibrational calculations, in which all nuclei are treated as quantum particles on a potential energy surface [4–8], the first step is the separation of the Cartesian coordinates of the center of mass followed by the definition of a body-fixed frame, orientational angles, and internal coordinates. This approach results in the replacement of the original laboratory-fixed Cartesian coordinates with curvilinear coordinates and the corresponding very complicated, translationally invariant rotational-vibrational Hamiltonians, see for example [9].

Less complicated translationally invariant Hamiltonians are used in full pre-BO calculations, where the original laboratory-fixed Cartesian coordinates are replaced by some set of translationally invariant Cartesian coordinates (TICCs) and the center-of-mass Cartesian coordinates (CMCCs) are separated [10–14]. Although the resulting TICCs are rectilinear coordinates, the corresponding Hamiltonian is still complicated. It is therefore reasonable to ask whether we can make our calculations even simpler, using the original laboratory-fixed-Cartesian-coordinate (LFCC) formalism without having to rely on any coordinate transformation at all. In this work we therefore explore the usage of LFCCs in pre-BO calculations.

If LFCCs are used one has to make sure that the energy of the overall translation of the system is eliminated. The most straightforward way is the subtraction of the kinetic energy operator of the center of mass from the total Hamiltonian, which, however, requires the evaluation of an additional integral with mixed coordinate second

derivatives [15, 16]. To avoid this additional integral evaluation we develop here an alternative approach.

In short, our computational strategy in the LFCC formalism to obtain eigenstates with various angular momentum quantum number is as follows. In our variational procedure we use basis functions, which are eigenfunctions of the total spatial angular momentum operators,  $\hat{L}^2$  and  $\hat{L}_z$  [14]. This is the simplest way to make sure that we obtain angular momentum eigenstates, since rotational “contamination” cannot be removed by a simple subtraction of a term from the full Hamiltonian [16, 17]. Then, we investigate the effect of the parametrisation of the basis functions on the translational contamination of the total energy and correct for it during the evaluation of the integrals in the LFCC formalism.

The paper is organised as follows. Firstly, in Section 2, we present the necessary theoretical details. In Section 3, we identify the translational energy contamination of the matrix elements. Next, in Section 4 we present numerical results for the two lowest rotational states of the  $\text{H}_2$  molecule.

## 2 Theoretical Details

In this section we present the theoretical details of the variational procedure applied to solve the full quantum Hamiltonian without clamping the nuclei. We introduce the notation essential to this work. Firstly, we define laboratory-fixed Cartesian coordinates (LFCCs) and translationally invariant Cartesian coordinates (TICCs) [18] as well as the corresponding non-relativistic quantum Hamiltonians. The introduction of the TICC formalism is necessary to better understand the mathematical properties of the LFCC formalism, which we then use in the calculations. Then, we define the basis functions using explicitly correlated Gaussian functions (ECGs) [19–23] and the global vector representation (GVR) [24–26]. Finally, we present the corresponding matrix elements and point out the parametrisation problems in the LFCC formalism.

The generalised eigenvalue problem corresponding to the matrix representation of the Hamiltonian is solved using the standard linear algebra library routines of LAPACK

(Version 3.2.1) [27] through the Armadillo framework (Version 3.4.0) [28].

## 2.1 Coordinates and the Quantum Hamiltonian

The non-relativistic quantum Hamiltonian for  $n + 1$  particles with  $m_i$  masses and  $q_i$  electric charges is in atomic units

$$\hat{H} = - \sum_{i=1}^{n+1} \frac{1}{2m_i} \Delta_{\mathbf{r}_i} + \sum_{i=1}^{n+1} \sum_{j>i}^{n+1} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \quad (1)$$

where the vector  $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{n+1})$  collects the LFCCs. The translational motion of the center of mass has traditionally been separated by introducing some set of TICC's [18],  $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  and the Cartesian coordinates of the center of mass (CMCC),  $\mathbf{X}_{\text{CM}}$ , by a linear transformation

$$\mathbf{x}_{\text{TICM}} = \begin{bmatrix} \mathbf{x} \\ \mathbf{X}_{\text{CM}} \end{bmatrix} = (\mathbf{U}_x \otimes \mathbf{1}_3) \mathbf{r} \quad (2)$$

where  $\mathbf{1}_3$  is the  $3 \times 3$  unit matrix and  $\mathbf{U}_x$  is a non-singular constant matrix with the restriction

$$\sum_{j=1}^{n+1} (\mathbf{U}_x)_{ij} = 0, \quad \text{with } i = 1, 2, \dots, n \quad (3)$$

and

$$(\mathbf{U}_x)_{n+1,j} = m_j / m_{\text{tot}}, \quad \text{with } j = 1, 2, \dots, n+1, \quad (4)$$

which guarantees the translational invariance for the coordinates  $\mathbf{x}$  and gives the definition of the center of mass. The notation  $m_{\text{tot}} = \sum_{j=1}^{n+1} m_j$  was introduced. The general form of the quantum Hamiltonian corresponding to the new coordinates  $\mathbf{x}$  and  $\mathbf{X}_{\text{CM}}$  [18, 26, 29] is

$$\hat{H}_{\text{TICM}} = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \mu_{ij}^{(x)} \nabla_{\mathbf{x}_i}^T \nabla_{\mathbf{x}_j} - \frac{1}{2m_{\text{tot}}} \Delta_{\mathbf{X}_{\text{CM}}} + \sum_{i=1}^{n+1} \sum_{j>i}^{n+1} \frac{q_i q_j}{|(\mathbf{f}_{ij}^{(x)} \otimes \mathbf{1}_3)^T \mathbf{x}|} \quad (5)$$

with  $\nabla_{\mathbf{x}_i}^T = (\partial/\partial x_{i1}, \partial/\partial x_{i2}, \partial/\partial x_{i3})$ ,

$$\mu_{ij}^{(x)} = \sum_{k=1}^{n+1} (\mathbf{U}_x)_{ik} (\mathbf{U}_x)_{jk} / m_k \quad \text{with } i, j = 1, 2, \dots, n \quad (6)$$

and

$$(\mathbf{f}_{ij}^{(x)})_k = (\mathbf{U}_x^{-1})_{ik} - (\mathbf{U}_x^{-1})_{jk} \quad \text{with} \quad i, j = 1, 2, \dots, n. \quad (7)$$

In Eq. (5) the translational kinetic energy of the center of mass,  $\hat{T}_{\text{CM}} = -1/2m_{\text{tot}}\Delta_{\mathbf{x}_{\text{CM}}}$ , separates, and by subtracting it the translationally invariant (TI) Hamiltonian,  $\hat{H}_{\text{TI}} = \hat{H} - \hat{H}_{\text{CM}}$ , is obtained. Mayer and Rokob [30] have recently shown that the same result can be obtained by requiring that the total momentum of the system is zero.

Traditionally the TI form of the Hamiltonian has been used in pre-BO calculations [10, 11, 14]. In this work we explore the consequences of the direct usage of the LFCC formalism and the original Hamiltonian, Eq. (1).

## 2.2 Basis Functions

The basis functions are defined similarly to Ref. [14]. The spatial part is constructed using ECGs [19–23] and the GVR [24, 25]. In what follows we introduce the notation relevant for this work.

The basic structure of a basis function written in LFCCs,  $\mathbf{r}$ , is

$$\psi_{LM_L}(\mathbf{r}; \mathbf{A}, \mathbf{u}, K) = |\mathbf{v}|^{2K+L} Y_{LM_L}(\hat{\mathbf{v}}) \exp\left(-\frac{1}{2}\mathbf{r}^T(\mathbf{A} \otimes \mathbf{1}_3)\mathbf{r}\right) \quad (8)$$

with the global vector

$$\mathbf{v} = (\mathbf{u} \otimes \mathbf{1}_3)\mathbf{r}. \quad (9)$$

The values collected in  $\mathbf{u} = (u_1, u_2, \dots, u_{n+1})$  are variational parameters. Further variational parameters are  $K \in \mathbb{N}_0$  and the elements of the symmetric matrix  $\mathbf{A}$  which describe correlations between the elements of  $\mathbf{r}$ . The matrix  $\mathbf{A}$  must be positive definite to ensure that the basis functions are square-integrable and have a non-vanishing norm.

Furthermore, in Eq. (8)  $Y_{LM_L}(\hat{\mathbf{v}})$  is a spherical harmonic function of degree  $L$  and order  $M_L$  and  $\hat{\mathbf{v}}$  represents the direction of the global vector. Thus, the parity of the basis function  $\psi_{LM_L}$  is  $(-1)^L$  (“natural parity”). The quantum numbers  $L$  and  $M_L$  correspond to the total “spatial” (orbital plus rotational) angular momentum operators,  $\hat{L}^2$  and  $\hat{L}_z$ .

Next, we consider the transformation of the LFCCs,  $\mathbf{r}$ , to some set of TICC,  $\mathbf{x}$  and the CMCCs,  $\mathbf{X}_{\text{CM}}$ , in order to better understand the properties of the LFCC formalism, which we intend to use in variational calculations. Upon a linear transformation of the coordinates the mathematical form of the spatial functions remains the same, and only the parameter matrices have to be transformed

$$\psi_{LM_L}(\mathbf{r}; \mathbf{A}, \mathbf{u}, K) = |\mathbf{v}|^{2K+L} Y_{LM_L}(\hat{\mathbf{v}}) \exp\left(-\frac{1}{2} \mathbf{r}^T (\mathbf{A} \otimes \mathbf{1}_3) \mathbf{r}\right) \quad (10)$$

$$= |\mathbf{v}|^{2K+L} Y_{LM_L}(\hat{\mathbf{v}}) \exp\left(-\frac{1}{2} \mathbf{x}_{\text{TICM}}^T (\mathbf{A}^{(x)} \otimes \mathbf{1}_3) \mathbf{x}_{\text{TICM}}\right) \quad (11)$$

with

$$\mathbf{v} = (\mathbf{u} \otimes \mathbf{1}_3) \mathbf{r} \quad (12)$$

$$= (\mathbf{u}^{(x)} \otimes \mathbf{1}_3) \mathbf{x}_{\text{TICM}} . \quad (13)$$

According to the definition of the  $\mathbf{x}_{\text{TICM}}$  coordinates, Eq. (2), the parameter matrices can be transformed back and forth as

$$\mathbf{A}^{(x)} = \mathbf{U}_x^{-T} \mathbf{A} \mathbf{U}_x^{-1} \quad \Leftrightarrow \quad \mathbf{A} = \mathbf{U}_x^T \mathbf{A}^{(x)} \mathbf{U}_x \quad (14)$$

and

$$\mathbf{u}^{(x)} = \mathbf{U}_x^{-T} \mathbf{u} \quad \Leftrightarrow \quad \mathbf{u} = \mathbf{U}_x^T \mathbf{u}^{(x)} \quad (15)$$

Due to the translational invariance condition, Eqs. (3)–(4), the parameter matrices have block structure [14]

$$\mathbf{A}^{(x)} = \begin{bmatrix} \mathcal{A}^{(x)} & 0 \\ 0 & c_A \end{bmatrix} \quad \text{and} \quad \mathbf{u}^{(x)} = \begin{bmatrix} \mathbf{u}^{(x)} \\ c_u \end{bmatrix} . \quad (16)$$

By combining this property with Eqs. (10)–(11) we obtain

$$\psi_{LM_L}(\mathbf{r}; \mathbf{A}, \mathbf{u}, K) = |\mathbf{v}|^{2K+L} Y_{LM_L}(\hat{\mathbf{v}}) \exp\left(-\frac{1}{2} \mathbf{r}^T (\mathbf{A} \otimes \mathbf{1}_3) \mathbf{r}\right) \quad (17)$$

$$= |\mathbf{v}|^{2K+L} Y_{LM_L}(\hat{\mathbf{v}}) \exp\left(-\frac{1}{2} \mathbf{x}^T (\mathcal{A}^{(x)} \otimes \mathbf{1}_3) \mathbf{x} - \frac{1}{2} c_A \mathbf{X}_{\text{CM}}^2\right) \quad (18)$$

and

$$\mathbf{v} = (\mathbf{u} \otimes \mathbf{1}_3) \mathbf{r} \quad (19)$$

$$= (\mathbf{u}^{(x)} \otimes \mathbf{1}_3) \mathbf{x} + c_u \mathbf{X}_{\text{CM}} , \quad (20)$$

which means that the system is at rest only in the case of  $c_A = 0$  and  $c_u = 0$  [14,31,32]. At the same time we have to ensure that our earlier requirements for the basis function parameters,  $\mathbf{u}$  and  $\mathbf{A}$  are fulfilled. While the condition  $c_u = 0$  does not cause any problems, we find that  $c_A = 0$  does so. If  $c_A = 0$  the matrix  $\mathbf{A}^{(x)}$  is singular and since  $\mathbf{A}$  and  $\mathbf{A}^{(x)}$  are related by a linear transformation, so is  $\mathbf{A}$ . This violates the requirement for  $\mathbf{A}$  to be positive definite. Thus, we can conclude that it is not possible to parametrise ECGs in the LFCC formalism so that the basis functions are square integrable with a non-vanishing norm and at the same time the system is at rest.

According to the block structure, Eq. (16) and the back-transformation, Eq. (14) the exponent matrix is constructed in our calculations according to

$$(\mathbf{A})_{ij} = -\alpha_{ij}(1 - \delta_{ij}) + \left( \sum_{k=1, k \neq i}^{n+1} \alpha_{ik} \right) \delta_{ij} + c_A \frac{m_i}{m_{\text{tot}}} \frac{m_j}{m_{\text{tot}}} . \quad (21)$$

## 2.3 Matrix Representation of the Hamiltonian

In spite of the parametrisation difficulties described in Section 2.2, we intend to use the LFCC formalism to construct the matrix representation for the quantum Hamiltonian because of its original simplicity. We repeat here only the necessary expressions from Ref. [14] and for the original integral derivation see [26].

The matrix element of the kinetic energy operator for the  $I$ th and  $J$ th quasi-normalised basis functions is [14]:

$$\begin{aligned} T_{IJ} &= -\frac{1}{2} \sum_{i=1}^n \frac{\langle \psi_I | \nabla_{\mathbf{r}_i}^T (\mathbf{M} \otimes \mathbf{1}_3) \nabla_{\mathbf{r}_i} | \psi_J \rangle}{|\psi_I| |\psi_J|} \\ &= D^{3/4} \left( \frac{p_{\mathbf{u}_I, \mathbf{u}_I}}{q_{\mathbf{u}_I}} \right)^{K_I} \left( \frac{p_{\mathbf{u}_J, \mathbf{u}_J}}{q_{\mathbf{u}_J}} \right)^{K_J} \left( \frac{p_{\mathbf{u}_I, \mathbf{u}_J}}{\sqrt{q_{\mathbf{u}_I} q_{\mathbf{u}_J}}} \right)^{L \min(K_I, K_J)} \sum_{m=0}^L \left( \frac{p_{\mathbf{u}_I, \mathbf{u}_J} p_{\mathbf{u}_I, \mathbf{u}_J}}{p_{\mathbf{u}_I, \mathbf{u}_I} p_{\mathbf{u}_J, \mathbf{u}_J}} \right)^m \\ &\times \left[ R_{IJ} + (K_I - m) \frac{P_{\mathbf{u}_I, \mathbf{u}_I}}{p_{\mathbf{u}_I, \mathbf{u}_I}} + (K_J - m) \frac{P_{\mathbf{u}_J, \mathbf{u}_J}}{p_{\mathbf{u}_J, \mathbf{u}_J}} + (L + 2m) \frac{P_{\mathbf{u}_I, \mathbf{u}_J}}{p_{\mathbf{u}_I, \mathbf{u}_J}} \right] H_{LK_I K_J m} \quad (22) \end{aligned}$$

where  $\mathbf{M}$  is a diagonal matrix with  $M_{ii} = 1/m_i$ . The  $H_{LK_I K_J m}$  terms are precalculated factors

$$H_{LK_I K_J m} = \frac{4^m (L + m + 1)!}{(K_I - m)! (K_J - m)! m! (2L + 2m + 1)!} \times \frac{1}{\sqrt{F_{K_I L} F_{K_J L}}} \quad (23)$$

in order to increase efficiency and ensure numerical stability with the terms

$$F_{KL} = \sum_{m=0}^K \frac{4^m (L+m+1)!}{(K-m)!(K-m)!m!(2L+2m+2)!} \quad (24)$$

stemming from the quasi-normalisation

$$|\psi_Z| = (\langle \psi_Z | \psi_Z \rangle)^{\frac{1}{2}} \quad \text{with} \quad Z \in \{I, J\}. \quad (25)$$

Furthermore we introduced short hand notations for terms which we have to study in terms of their dependence on  $c_A$ . One term we have to study is

$$D = \frac{\det(2\mathbf{A}_I) \det(2\mathbf{A}_J)}{\det(\mathbf{A}_{IJ}) \det(\mathbf{A}_{IJ})} \quad (26)$$

stemming from the origin centered Gaussian functions. Furthermore we have to analyse the terms depending on the global vectors

$$p\mathbf{u}_X, \mathbf{u}_Y = \mathbf{u}_X^T \mathbf{A}_{IJ}^{-1} \mathbf{u}_Y \quad \text{with} \quad X, Y \in \{I, J\} \quad (27)$$

$$P\mathbf{u}_I, \mathbf{u}_I = -\mathbf{u}_I^T \mathbf{A}_{IJ}^{-1} \mathbf{A}_J \mathbf{M} \mathbf{A}_J \mathbf{A}_{IJ}^{-1} \mathbf{u}_I \quad (28)$$

$$P\mathbf{u}_J, \mathbf{u}_J = -\mathbf{u}_J^T \mathbf{A}_{IJ}^{-1} \mathbf{A}_I \mathbf{M} \mathbf{A}_I \mathbf{A}_{IJ}^{-1} \mathbf{u}_J \quad (29)$$

$$P\mathbf{u}_I, \mathbf{u}_J = \mathbf{u}_I^T \mathbf{A}_{IJ}^{-1} \mathbf{A}_J \mathbf{M} \mathbf{A}_I \mathbf{A}_{IJ}^{-1} \mathbf{u}_J, \quad (30)$$

the terms from the quasi-normalisation

$$q\mathbf{u}_Z = \frac{1}{2} \mathbf{u}_Z^T \mathbf{A}_Z^{-1} \mathbf{u}_Z \quad \text{with} \quad Z \in \{I, J\} \quad (31)$$

and the term

$$R_{IJ} = \frac{3}{2} \text{Tr} [\mathbf{A}_{IJ}^{-1} \mathbf{A}_J \mathbf{M} \mathbf{A}_I] \quad (32)$$

which can be associated with the radial motion of the system [33].

Here, we immediately see that the singularity of  $\mathbf{A}_I$  and  $\mathbf{A}_J$  would cause terms in Eqs. (26)–(32) to be not defined. The singularity is introduced if the  $c_A = 0$  selection is made to guarantee translation free expressions.

Since the potential energy terms in Eq. (1) depend only on the inter-particle distances, we can choose any  $c_A > 0$  value and evaluate the matrix elements without any problem and the resulting potential energy matrix elements are independent of the value of  $c_A$ . So, they are not discussed here any longer and are evaluated according to Ref. [14].



### 3 Identification and Elimination Strategy for the Translational Contamination

In Section 2.2, we noted that the ECGs take the same mathematical form, independent of whether we choose the LFCCs  $\mathbf{r}$ , or some  $\mathbf{x}_{\text{TICM}}$ . This simple transformation behavior also transfers to the expressions of the integrals. Due to this property we can study the influence of  $c_A$  on the different terms in Eqs. (26)–(32), only the parameter matrices  $\mathbf{A}_I, \mathbf{A}_J$  and  $\mathbf{u}_I, \mathbf{u}_J$  corresponding to  $\mathbf{r}$  have to be replaced by their according expressions in terms of  $\mathbf{A}_I^{(x)}, \mathbf{A}_J^{(x)}$  and  $\mathbf{u}_I^{(x)}, \mathbf{u}_J^{(x)}$  corresponding to  $\mathbf{x}_{\text{TICM}}$ . The parameter matrices are related by the transformation given in Eqs. (14)–(15), and  $\mathbf{A}_I^{(x)}, \mathbf{A}_J^{(x)}$  and  $\mathbf{u}_I^{(x)}, \mathbf{u}_J^{(x)}$  have block structure, Eq. (16).

Firstly, let us consider the  $R_{IJ}$  term of the kinetic energy matrix elements, Eq. (32), explicitly. We analyze the properties of  $R_{IJ}$  with respect to  $c_A$  using the TICC formalism:

$$\begin{aligned} R_{IJ} &= \frac{3}{2} \text{Tr} [\mathbf{A}_{IJ}^{-1} \mathbf{A}_J \mathbf{M} \mathbf{A}_I] \\ &= \frac{3}{2} \text{Tr} [(\mathbf{A}_{IJ}^{(x)})^{-1} \mathbf{A}_J^{(x)} \mathbf{U}_x \mathbf{M} \mathbf{U}_x^T \mathbf{A}_I^{(x)}]. \end{aligned} \quad (33)$$

and also exploit the block structure of the matrices:

$$\mathbf{A}_z^{(x)} = \begin{bmatrix} \mathcal{A}_z^{(x)} & 0 \\ 0 & c_A \end{bmatrix} \quad \text{with } z \in \{I, J, IJ\} \quad \text{and} \quad \mathbf{U}_x \mathbf{M} \mathbf{U}_x^T = \begin{bmatrix} \boldsymbol{\mu}^{(x)} & 0 \\ 0 & c_M \end{bmatrix}. \quad (34)$$

The block structure of  $\mathbf{U}_x \mathbf{M} \mathbf{U}_x^T$  follows from the conditions of translation invariance, Eqs. (3)–(4), for  $\mathbf{x}$ . We can thus separate the  $c_A$  dependent terms in  $R_{IJ}$ :

$$R_{IJ} = R_{IJ}^{\text{Int}} + \frac{3}{4} c_A c_M = \frac{3}{2} \text{Tr} [(\mathcal{A}_{IJ}^{(x)})^{-1} \mathcal{A}_J^{(x)} \boldsymbol{\mu}^{(x)} \mathcal{A}_I^{(x)}] + \frac{3}{4} c_A c_M \quad (35)$$

The  $c_M$  factor of the linear contribution is

$$c_M = (\mathbf{U}_x \mathbf{M} \mathbf{U}_x^T)_{n+1, n+1} = \sum_{i=1}^{n+1} (\mathbf{U}_x)_{n+1, i}^2 / m_i = 1 / m_{\text{tot}}. \quad (36)$$

and thus with  $c_M = 1 / m_{\text{tot}}$

$$R_{IJ} = R_{IJ}^{\text{Int}} + \frac{3}{4} \frac{c_A}{m_{\text{tot}}} \quad (37)$$

Next, we investigate the  $c_A$  dependence of the remaining terms and factors, Eqs. (26)–(31), of the matrix representation of the kinetic energy in the LFCC formalism, Eq. (22). In Eq. (26) we find that any contribution of  $c_A$  cancels. In Eqs. (31) and (27) we see that  $c_A$  only contributes if  $c_u > 0$ . Since we require  $c_u = 0$ , this contribution is eliminated. Finally, in Eqs. (28)–(30) we find that the contribution of  $c_A$  to the exponent matrices cancel. In short, only the  $R_{IJ}$  term, Eq. (32), has a non-vanishing (linear)  $c_A$  dependence in the kinetic energy matrix element,  $T_{IJ}$ .

Thus, if  $c_A = 0$  was chosen, the translational dependence vanishes, but the exponents matrices,  $\mathbf{A}_I, \mathbf{A}_J$ , are singular without having an inverse. Thus, for an implementation in a computer program we can choose a non-zero value for  $c_A$ , and eliminate the translational contribution explicitly by subtracting  $3c_A/(4m_{\text{tot}})$  from the  $R_{IJ}$  matrix element. This is a simple computational strategy which we are going to follow in the LFCC formalism.

We also note here that the direct variational optimisation of all parameters, including  $c_A$  here, would be another option that has been suggested already in the literature, e.g., in Ref. [2]. From the theoretical details presented so far, we understand that the total energy with  $c_A > 0$  is always an upper bound to the total energy free of the overall translation of the system,  $E_{\text{tot}}(c_A) \geq E_{\text{TI}}$ . But as we have shown for  $c_A = 0$  several parameter matrices incorporated in  $E_{\text{tot}}(c_A)$  are singular, so in spite of the fact that the limit exists, the application in a computer program is problematic ( $c_u = 0$  is chosen throughout the discussion).

This explains our preference for the approach developed here, which releases the translation-free condition for the basis functions and corrects for the translational contamination in the kinetic energy explicitly for each basis function,  $I = 1, 2, \dots, N$ . The details of our algorithm are:

1. Generate, optimise or read in the  $\alpha_{I,ij}$  values for  $i = 1, 2, \dots, n+1$ ,  $j = i+1, i+2, \dots, n+1$ .
2. Construct the elements of the exponent matrix in the LFCC formalism as

$$(\mathbf{A}_I)_{ij} = -\alpha_{I,ij}(1 - \delta_{ij}) + \left( \sum_{k=1, k \neq i}^{n+1} \alpha_{I,ik} \right) \delta_{ij} + c_A \frac{m_i}{m_{\text{tot}}} \frac{m_i}{m_{\text{tot}}}$$

with  $i, j = 1, 2, \dots, n+1$  and  $c_A > 0$ .

3. Due to the  $c_A > 0$  choice the matrix  $\mathbf{A}_I$  is non-singular, and  $1/\det(\mathbf{A}_I)$  and  $\mathbf{A}_I^{-1}$  can be calculated. At the same time the total kinetic energy contains some translational contamination.
4. The translational contamination is eliminated by replacing  $R_{IJ}$ , Eq. (32), with  $R_{IJ} - 3c_A/(4m_{\text{tot}})$  in the expression of the kinetic energy matrix element,  $T_{IJ}$ , Eq. (22).

Throughout this computational strategy for the elimination of the translational contamination in the LFCC formalism we have  $c_u = 0$ ,  $c_A > 0$ .

## 4 Numerical Examples

In this section we present numerical applications using the LFCC formalism. The appearance of the translational contamination and its elimination according to the strategy described in Section 3 are demonstrated.

Our test cases are the lowest energy levels of the para-H<sub>2</sub> ( $L = 0, p = +1$ ) and the ortho-H<sub>2</sub> ( $L = 1, p = -1$ ) molecules both in the singlet electronic state. These are the two lowest-energy rotational states of the hydrogen molecule.  $L$  is the total spatial (orbital and rotational) quantum number and  $p$  is the parity. The wave functions are obtained by a direct solution of the linear variational problem using 1500 basis functions with an optimised parametrisation taken from [14]. Here we use the LFCC formalism exclusively, thus all parameters were transformed first to the LFCC representation, according to Eqs. (14)–(15). The exponent matrix for each basis function was constructed according to Eq. (21). During this transformation we were free to choose  $c_A$  to simulate different levels of translational contamination, while  $c_u = 0$  was fixed. We used the same  $c_A$  value for each basis function. Table I collects the results of our numerical calculations.

[Table 1 about here.]

The first column of Table I lists the five values of  $c_A$  which have been studied. The next two columns provide the total energies of the system with and without translational contamination, respectively. The last two columns list the kinetic energy contribution with and without translational contamination, respectively. We observe that those energies, which contain translational contributions (columns 2 and 4) depend on the value of  $c_A$  and that an increase of the value of  $c_A$  causes an increase of the energy, according to Section 3. We also note that the corrected energies (columns 3 and 5) are independent of  $c_A$ . Furthermore, we list the translational correction given in Eq. (37). Hence, these results give a numerical confirmation that we have identified and eliminated the translational contamination depending on the  $c_A$  basis function parameter, while choosing  $c_u = 0$ .

## 5 Conclusion

In this work, we used for the first time laboratory-fixed Cartesian coordinates (LFCCs) with an explicit correction strategy for the translational contamination in pre-Born–Oppenheimer variational calculations with various angular momentum quantum numbers. Our work was motivated by the inherent simplicity of the LFCCs and the corresponding operators.

Instead of transforming the coordinates we accounted for the translational and rotational invariances of the isolated many-particle problem by using an appropriate form and parametrisation of the basis functions in the variational procedure. The basis functions were constructed using explicitly correlated Gaussians and the global vector representation, and as an extension of our earlier work [14] we focused here on the usage of LFCCs and the problem of translational invariance.

First of all, we observed that it is impossible to parametrise explicitly correlated Gaussian functions (ECGs) in such a way that the total system is at rest in LFCCs and at the same time the basis functions are square-integrable with a non-vanishing norm.

Fortunately, it was possible to devise a simple computational strategy to circumvent this problem. So, for the sake of the stability of the numerical computations we released the translational constraint on the basis functions, by choosing a non-zero value for the free parameter,  $c_A$ , in the LFCC parametrisation of the ECG exponents, and then explicitly corrected for the translational contamination in the kinetic energy integral expressions. This correction term is a simple constant, which depends linearly on  $c_A$ . Its form was derived by considering a few mathematical relationships of the formalism: a) the properties of the linear transformation between LFCCs and translationally invariant and center-of-mass Cartesian coordinates (TICMCCs); b) the corresponding transformation of the basis function parameter matrices; c) the fact that the parameter matrices are block diagonal in TICMCCs.

It was also shown that the uncorrected total energy  $E_{\text{tot}}(c_A)$  is an upper bound to the translation-free total (intrinsic) energy. Thus, in principle, we could obtain this value by the variational optimisation of the LFCC parametrisation (implicitly including the  $c_A$  value in the optimisation). We prefer, however, our explicit treatment for the elimination of the translational contamination, because in the  $c_A = 0$  limit the exponent matrix,  $\mathbf{A}$ , of each ECG would be singular, and thus the numerical evaluation of  $1/\det(\mathbf{A})$  and  $\mathbf{A}^{-1}$  would be impossible.

Finally, to demonstrate the numerical applicability of our approach we calculated the lowest two rotational levels of the singlet hydrogen molecule corresponding to the para and ortho proton spin states, respectively.

The presented LFCC formalism with the explicit translational contamination correction is an alternative but equivalent to the traditional approaches using some set of translationally invariant Cartesian coordinates with the Cartesian coordinates of the center of mass explicitly separated already in the Hamiltonian, e.g. [10, 14]. The simplicity of the LFCCs is an appealing choice for the variational solution of the Schrödinger equation with the non-relativistic Hamiltonian. Furthermore, one can think of more complicated operators for which the usage of the simplest possible coordinate representation and the avoidance of any coordinate transformation is more than just a comfortable option.

## 6 Acknowledgments

This work has been supported by the Swiss National Science Foundation SNF (project 200020\_144458/1). EM. thanks the Hungarian Scientific Research Fund (OTKA, NK83583) for financial support.

## References

- [1] A. Chakraborty, M.V. Pak and S. Hammes-Schiffer, J. Chem. Phys. **129** (1), 014101 (2008).
- [2] A.D. Bochevarov, E.F. Valeev and C.D. Sherrill, Mol. Phys. **102** (1), 111 (2004).
- [3] M. Goli and S. Shahbazian, Theor. Chem. Acc. **131** (5), 1208 (2012).
- [4] S. Albert, H. Hollenstein, M. Quack and M. Willeke, Mol. Phys. **104** (16-17), 2719 (2006).
- [5] S. Albert, K.K. Albert, H. Hollenstein, C.M. Tanner and M. Quack, Fundamentals of RotationVibration Spectra, in *High-Resolution Spectroscopy*, edited by Martin Quack and Frederic Merkt (John Wiley and Sons, Ltd., Chichester, 2011).
- [6] D. Luckhaus, M. Quack and M. Willeke, Z. Physik. Chem. **1114** (8), 1087 (2000).
- [7] B. Kuhn, T.R. Rizzo, D. Luckhaus, M. Quack and M.A. Suhm, J. Chem. Phys. **111** (6), 2565 (1999).
- [8] M. Quack, Chimia **55** (10), 753 (2001).
- [9] E. Mátyus, G. Czakó and A.G. Császár, J. Chem. Phys. **130** (13), 134112 (2009).
- [10] M. Cafiero, S. Bubín and L. Adamowicz, Phys. Chem. Chem. Phys. **5** (8), 1491 (2003).
- [11] S. Bubín, M. Pavanello, W.C. Tung, K.L. Sharkey and L. Adamowicz, Born-Oppenheimer and Non-Born-Oppenheimer, Atomic and Molecular Calculations with Explicitly Correlated Gaussians. 2012, (in press) (unpublished), <<http://www.ncbi.nlm.nih.gov/pubmed/23020161>>.
- [12] E. Mátyus, J. Hutter, U. Müller-Herold and M. Reiher, Phys. Rev. A **83** (5), 052512 (2011).

- [13] E. Mátyus, J. Hutter, U. Müller-Herold and M. Reiher, J. Chem. Phys. **135** (20), 204302 (2011).
- [14] E. Mátyus and M. Reiher, J. Chem. Phys. **137** (2), 024104 (2012).
- [15] H. Nakai, Int. J. Quantum Chem. **86** (6), 511 (2002).
- [16] H. Nakai, M. Hoshino, K. Miyamoto and S. Hyodo, J. Chem. Phys. **122** (16), 164101 (2005).
- [17] B. Sutcliffe, J. Chem. Phys. **123** (23), 237101 (2005).
- [18] B.T. Sutcliffe, Chapter 31 Coordinate Systems and Transformations, in *Handbook of Molecular Physics and Quantum Chemistry*, edited by Stephen Wilson (John Wiley & Sons, Ltd, Chichester, 2003), pp. 485–500.
- [19] B. Jeziorski and K. Szalewicz, Phys. Rev. A **19** (6), 2360 (1979).
- [20] W. Cencek and J. Rychlewski, J. Chem. Phys. **98** (2), 1252 (1993).
- [21] J. Rychlewski, editor, *Explicitly Correlated Wave Functions in Chemistry and Physics* (, , 2003).
- [22] S.F. Boys, Proc. R. Soc. London Ser. A **258** (1294), 402 (1960).
- [23] K. Singer, Proc. R. Soc. London Ser. A **258** (1294), 412 (1960).
- [24] K. Varga, Y. Suzuki and J. Usukura, Few-Body Syst. **24** (1), 81 (1998).
- [25] Y. Suzuki, J. Usukura and K. Varga, J. Phys. B **31** (1), 31 (1998).
- [26] Y. Suzuki and K. Varga, *Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems, Lect. Notes Phys.*, Vol. 54 (Springer, Berlin, Heidelberg, 1998).
- [27] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen, *LAPACK Users' Guide*, 3rd ed. (Society for Industrial and Applied Mathematic, Philadelphia, PA, 1999).
- [28] C. Sanderson, Armadillo: An open source C++ linear algebra library for fast prototyping and computationally intensive experiments, , NICTA 2010 .
- [29] B.T. Sutcliffe, Chapter 32 Molecular Hamiltonian, in *Handbook of Molecular Physics and Quantum Chemistry*, edited by Stephen Wilson, Vol. 1 (John Wiley & Sons, Ltd, Chichester, 2003), pp. 501–525.

- [30] I. Mayer and T. Rokob, Phys. Rev. A **85** (4), 044101 (2012).
- [31] G. Hagedorn, Comm. Math. Phys. **93**, 77 (1980).
- [32] E. Faou, V. Gradinaru and C. Lubich, SIAM J. Sci. Comput. **31** (4), 3027 (2009).
- [33] D.B. Kinghorn, Int. J. Quantum Chem. **57** (2), 141 (1996).



# List of Tables

I	The appearance and elimination of the translational contamination of the total pre-Born–Oppenheimer energy using the laboratory-fixed-Cartesian-coordinate (LFCC) formalism. The ground-state energies, in $E_h$ , of the singlet, para- $H_2$ ( $L = 0, p = +1$ ) and ortho- $H_2$ ( $L = 1, p = -1$ ) molecules are calculated using the basis function parameter set of Ref. [14]. The corresponding translationally invariant energies are $E_{TI}(L = 0) = -1.164\,025\,026\,E_h$ and $E_{TI}(L = 1) = 1.163\,485\,167\,E_h$ , respectively. <sup>a</sup> Results corrected for the translational contamination of the LFCC formalism are indicated with the "corr." subscript. . . . .	18
---	--	----

Table I: The appearance and elimination of the translational contamination of the total pre-Born–Oppenheimer energy using the laboratory-fixed-Cartesian-coordinate (LFCC) formalism. The ground-state energies, in  $E_h$ , of the singlet, para- $H_2$  ( $L = 0, p = +1$ ) and ortho- $H_2$  ( $L = 1, p = -1$ ) molecules are calculated using the basis function parameter set of Ref. [14]. The corresponding translationally invariant energies are  $E_{TI}(L = 0) = -1.164025026 E_h$  and  $E_{TI}(L = 1) = 1.163485167 E_h$ , respectively.<sup>a</sup> Results corrected for the translational contamination of the LFCC formalism are indicated with the ”corr.” subscript.

$c_A^b$	$E_{LF}^c$	$E_{LF,corr.}^c$	$\langle \Psi_{LF}   \hat{T}_{LF}   \Psi_{LF} \rangle^c$	$\langle \Psi_{LF}   \hat{T}_{LF}   \Psi_{LF} \rangle_{corr.}^c$	$\delta_{Tr.}^c$
para- $H_2$ $L = 0$					
0.01	-1.164022985	-1.164025026	1.164027045	1.164025004	0.000002041
0.10	-1.164004614	-1.164025026	1.164045416	1.164025004	0.000020412
0.50	-1.163922966	-1.164025026	1.164127064	1.164025004	0.000102060
1.00	-1.163820906	-1.164025026	1.164229124	1.164025004	0.000204120
2.00	-1.163616786	-1.164025026	1.164433245	1.164025004	0.000408240
ortho- $H_2$ $L = 1$					
0.01	-1.163483125	-1.163485167	1.163487203	1.163485161	0.000002041
0.10	-1.163464755	-1.163485167	1.163505573	1.163485161	0.000020412
0.50	-1.163383107	-1.163485167	1.163587222	1.163485161	0.000102060
1.00	-1.163281047	-1.163485167	1.163689282	1.163485161	0.000204120
2.00	-1.163076927	-1.163485167	1.163893402	1.163485161	0.000408240

<sup>a</sup> The parametrisation of the internal basis set ( $\mathcal{P}$ ) was taken from Ref. [14]. The number of basis functions was 1 500, the order of the polynomial prefactors was  $2K \in [0, 20]$ . For the proton-electron mass ratio the value  $m_p/m_e = 1836.15267247$  was used, and thus  $m_{tot} = 3674.30534494 m_e$ .

<sup>b</sup> Free parameter of the explicitly correlated Gaussian basis functions expressed in LFCCs. The same  $c_A$  value was used for each basis function in the basis set. The value  $c_u = 0$  was used throughout the calculations.

<sup>c</sup>  $E_{LF}$ ,  $\Psi_{LF}$  : Eigenvalue and eigenfunction of the full Hamiltonian expressed in LFCC using the basis function parameters ( $\mathcal{P}, c_A$ ).

$\hat{T}_{LF}$  : The kinetic energy operator expressed in LFCCs.

”corr.” : Correction for the translational contamination in the LFCC formalism, as explained in Section 2.

$\delta_{Tr.} = 3c_A/(4m_{tot})$ : Translational contamination of the kinetic energy, see Eq. (37). The corrected, translation-free value is obtained as  $E_{LF,corr.} = E_{LF} - \delta_{Tr.}$ .